


Theresa Rogers  
Kari Lynn Winters  
Gregory Bryan  
John Price  
Frank McCormick  
Liisa House  
Dianna Mezzarobba  
Carollyne Sinclair

# Developing the IRIS: Toward situated and valid assessment measures in collaborative professional development and school reform in literacy



*The IRIS is an assessment tool developed in the context of a literacy reform effort.*

In the context of a collaborative literacy reform effort, we explored the use of a situated assessment tool to support and evaluate our project, providing an alternative or complement to standardized or published tests. By situated assessment, we mean an assessment tool that is collaboratively developed and used in the context of a particular reform effort and is meant to benefit teachers who use it by informing their instruction. We contrast this to standardized measures that are used solely for accountability rather than for professional development (cf. Dillon, 2003).

The role of the first three authors (university researchers Rogers, Winters, and Bryan) was to help conceptualize, support, and evaluate the project along with five teachers from the local school board who serve as consultants and mentor teachers (authors Price, McCormick, House, Mezzarobba, and Sinclair). This project required us to build relationships among administrators, teachers, and researchers to assess progress and pitfalls and respond to the gaps and tensions in the

project through collaborative “on-the-ground” theorizing—particularly about the uses of assessment in school reform.

## The context

Many educational reform initiatives in North America include external assessments from state or provincial governments, or from publishing companies, to make decisions about the nature and outcome of the reform efforts. However, in contrast to the current climate of reform in the United States, where accountability is tied to high-stakes standardized testing, provincial governments in Canada have asked individual districts to create their own accountability contracts with a range of possible indicators of success. The school board (district) with which we worked has named literacy as the key area for improvement in the current accountability contract, and both provincial tests and school or classroom-based measures are included as indicators of progress.

Within the context of this literacy reform effort in western Canada, we worked together to develop a project to improve reading comprehension

among students in grades 4 through 8. The call for increased comprehension instruction is hardly new (cf. Durkin, 1978/1979); however, there is evidence to suggest that comprehension strategy instruction is still not getting the attention it deserves as a critically important component of effective and successful reading programs (Taylor, Pearson, Clark, & Walpole, 2002). Indeed, Cassidy and Cassidy (2004/2005) have cited comprehension as a “hot topic.” This should not be surprising given that reading text strategically is particularly important in a time when the ability to garner, negotiate, synthesize, and critique information across a range of print and nonprint genres is acknowledged as a key component of participating in an information or knowledge economy (cf. Luke & Elkins, 1998).

On the basis of our evaluation of the first year of the project, we decided that collaboratively developing a situated and valid reading strategies measure would be key to supporting effective professional development in the larger school reform project. A major goal of this literacy project was to improve reading comprehension by supporting students from grades 4 to 8 in the development and use of comprehension strategies (Pearson & Dole, 1987) in flexible ways across a range of texts, with a focus on informational texts. We developed the project around six reading comprehension strategies: Making Connections, Engaging With the Text, Active Meaning Construction, Monitoring Understanding, Analysis and Synthesis, and Critical Reading (see Table 1).

These comprehension strategies are adapted from the work of Pearson and colleagues (Duke & Pearson, 2002; Fielding & Pearson, 1994; Pearson, Roehler, Dole, & Duffy, 1992) and are compatible, for instance, with the best practice comprehension methods recommended by the National Reading Panel (National Institute of Child Health and Human Development, 2000).

## Background

Results from a provincial-wide fundamental skills assessments in reading were one impetus for the school board to develop the comprehension focus of the literacy project: In 2001, only 77% of grade 4 students in the district met expectations in

**TABLE 1**  
**A description of the strategies on the IRIS**

- **Making Connections:** drawing on background knowledge and experiences, using your knowledge of genre and the author’s craft, developing purposes for reading (e.g., what you would like to learn).
- **Engaging With the Text:** visualizing, entering the world of the text, taking perspectives, responding to the author’s craft.
- **Active Meaning Construction:** questioning, connecting ideas, hypothesizing and drawing inferences, predicting, and using text-structure cues.
- **Monitoring Understanding:** awareness of lack of understanding, cross-checking ideas, revising hypotheses and predictions, using repair/fix-up strategies.
- **Analysis and Synthesis:** distinguishing more important ideas from less important ones, drawing conclusions from the ideas in the text, constructing themes, connecting information across texts (and other media).
- **Critical Reading:** questioning the author, the relationship of the text to yourself and to the world, evaluating ideas, developing alternative interpretations.

reading comprehension; at grade 7, it was only 73%; and at grade 10, it was only 74%. At the same time, a grade 8 (the first year of high school in this province) formal assessment of incoming students indicated that, at some schools, a considerable proportion of the student population was reading below grade level. In addition, anecdotal accounts by teachers repeatedly indicated that significant numbers of their students struggled to comprehend texts across the curriculum. These were among the reasons the board initiated the 4 to 8 Literacy Project as a major professional development and reform effort and included it in the 2003/2004 District Accountability Contract, which cites the improvement of literacy as a main goal.

While naming literacy as a priority, we recognized that successful school reform is closely related to effective teacher professional development (Fullan, 1992). Indeed, professional development is often recognized as one of the key ingredients to which the successful implementation of school reform might be attributed (e.g., Taylor et al., 2002). While many teachers develop effective techniques in one aspect or another of the literacy curriculum,

they also need time to share and build on that acquired knowledge and experience and to connect it to the growing body of literature in this area. To effect real improvement in schools, teachers need an opportunity to engage in sophisticated and long-term professional support that allows for inquiry, reflection, and dialogue and that provides instructional interventions to meet specific goals. These interventions also need to be relatively easy to implement and adapt to particular classroom settings (cf. Anders, Hoffman, & Duffy, 2000).

The 4 to 8 Literacy Project drew on models of school improvement that include focused professional development work by classroom teachers who share beliefs and understandings about their goals (Hill, 1998); an emphasis on a range of teaching approaches and integration of assessment and instruction (Langer, 1999); and an opportunity for teachers to exchange ideas and gain new knowledge, reflect on their current and new practices, and provide feedback on the goals and methods of the project. As we (the project team) reflected on the professional development of the first year, however, we realized that a missing aspect of the project was a reading comprehension strategies assessment tool that might connect teachers' understanding of the comprehension strategies with their professional development goals.

Our evaluation of the first year of the project included an analysis of school board documents, teacher surveys from the beginning and end of the year, and examinations of field notes from teacher-leader meetings for the project and from meetings at one case study school. We found that the implementation of an informal process writing assessment in the first year was key to the teachers' engagement in and understanding of their students' strengths and weaknesses in writing. At that point, the reading component of the project was being assessed with a standardized test (Canadian Achievement Test, or CAT). This test, however, was not yielding similar engagement by the teachers in understanding students' reading abilities and the role of teaching reading strategies in the curriculum. We recognized the need for a reading comprehension strategies measure that was less formal and addressed the needs of the teachers and students in the project. We then undertook the collaborative development of the Informal Reading

Inventory of Strategies (IRIS) to support and enhance the project goals.

## Grappling with issues of validity

As we began to think about developing this tool, we recognized the need to think not only about issues of content validity (measuring use of the comprehension strategies that were at the core of the project), but also to think about the potential uses of the measure at all levels of the project (see Table 2).

The concept of validity in testing is still being established (Hubley & Zumbo, 1996) and has expanded beyond notions of content validity tied to issues of procedure to a more dynamic notion tied to social processes and interpretations (Murphy, 1998). For instance, the International Reading Association and the National Council of Teachers of English (1994) Joint Task Force on Assessment called for measures that, among other standards, have as their purpose the improvement of teaching and learning, recognize and reflect the intellectually and socially complex nature of reading and writing in various contexts, take into account the interests of the students, are fair and equitable, consider the consequences of assessment, and allow for critical inquiry into curriculum and instruction.

Definitions such as these include more expansive views of validity that incorporate not only questions of what is being tested but also who is being tested, why and how they are being tested, as well as the consequences of testing (Johnston, 1998; Moss, 1998); that is, validity in measurement and literacy "is a complex interplay between evidence and values" (Murphy, 1998, p. 27).

In collaboratively developing the IRIS, we took into account these notions of consequential validity (Johnston, 1998), including the idea that development of assessment measures should take into account the instructional effects in the educational context (cf. Frederiksen & Collins, 1989). Finally, we were also influenced by the argument that validity should be responsive and transactional (Tierney, Crumpler, Bertelsen, & Bond, 2003) by including stakeholders in the negotiation and development of the assessment measures.

## The structure of the project

The university involvement in the project was part of a larger school board and university partnership with literacy as one major focus along with technology, social development, Indigenous education, and research and development. The school board serves approximately 70,000 students and has a diverse population in terms of ethnicity and language. This three-year literacy project began in the 2002–2003 school year, with approximately 100 teachers and 2,500 students participating. The focus of the project was both reading comprehension and writing, but the main focus of the university partnership was on supporting and evaluating the reading comprehension component.

This framework for teaching reading comprehension was adopted by the 12 schools (7 elementary and 5 secondary) encompassing teachers in grades 4 through 8 who volunteered to participate in the project. Support was provided to over 100 participating teachers and approximately 25 teacher leaders through mentorship, professional development, and material resources.

The project team members included three university researchers, two school board curriculum consultants (intermediate and secondary) with over 20 years' experience in elementary and secondary classrooms, two literacy mentors who are part-time teachers, and one Aboriginal support consultant. Each school also has one to three project leaders and up to eight teachers of grades 4 through 8 involved in the project.

The professional development and support is extensive and varied in the project. There were team meetings to conceptualize the project and plan professional development. Teacher-leader meetings were held several times a year, and schools were given resources to participate in various sorts of professional development workshops, all-day project conferences, and school-team professional development time along with time to visit other classrooms. The two mentor teachers and the two school board consultants provided professional development workshops on reading strategies as requested. Each school is also provided with a “resource tub” of professional books and teaching materials.

The role as university consultants and collaborators was complex because we have helped to conceptualize the project, support its implementation,

**TABLE 2**  
**Validity in assessment**

*Content validity:* Assumes the knowledge and skills assessed represent the larger domain of knowledge and skills that are being taught (in this case, reading comprehension strategies).

*Consequential validity:* Takes into account how and why the assessment is used and who is being assessed.

*Transactional validity:* Involves all stakeholders in the negotiation and development of the assessment measures, including teachers, students, and administrators.

participate in professional development as speakers and workshop leaders, and plan evaluation. However, these multiple levels of involvement contributed to our approach to developing a more valid and situated reading comprehension strategies measure.

## Developing and implementing the IRIS

In response to the need for an assessment tool for this project, initially we attempted to develop rubrics for the teachers to use with any text or lesson they chose. We developed several rubric prototypes with the project team and shared them at school meetings. The feedback from teachers made it clear that they preferred an assessment tool that would be complete with a passage, questions, and scoring rubric to measure comprehension strategy use among their students. It was also evident from conversations with teachers that they wanted the scoring rubric to correspond to their provincial literacy rubrics outlined in a document on reading performance standards from the Ministry of Education that used a four-level scale (1 = not yet meeting expectations, 2 = minimally meeting expectations, 3 = fully meeting expectations, and 4 = exceeding expectations). The corresponding scale developed for the IRIS is 1 = no use of strategies, 2 = minimal use of strategies, 3 = appropriate use of strategies, and 4 = rich use of strategies.

In our visits to project school sites, teachers also shared with us other reading assessment tools

they had seen, including one from a local district that included an oral reading measure that they thought would be useful for grouping students early in the year. We decided to include an oral reading excerpt and to provide the scoring guide for miscue analysis (adapted from Rhodes & Shanklin, 1990). The most important aspect of this assessment was its potential to be used by teachers to guide their instructional decision making and practice; that is, that it carried validity in terms of use as well as content.

### **The IRIS written form**

We wanted the assessment tool to be as transparent as possible to students and teachers, so we listed the reading strategies that are central to the project on the first page of the written or student form (see Table 1). Below that, we listed two pre-reading questions focusing on connecting ideas (see questions 1 and 2 in Table 3). After reading the first part of the nonfiction selection, students answer several more questions focusing on constructing meaning and engaging with text (see questions

**TABLE 3**  
**Pre-, during-, and postreading questions on the IRIS**

1. After taking a quick look at the reading selection, do you think this passage is fiction (e.g., a story) or nonfiction (true)? What clues helped you decide? (Making Connections)
2. What do you already know about this topic? (Making Connections)
3. Now that you have read the first part, what do you think you will find out in the rest of the passage? (Active Meaning Construction)
4. If you drew a picture of something in this passage, what would it look like? (Engaging With the Text)
5. What do you find most interesting so far? (Engaging With the Text)
6. What do you think were the two or three most important ideas in the passage? (Analysis and Synthesis)
7. If the author were sitting here, what questions would you ask him or her? (Critical Reading)
8. If someone were having problems understanding this passage, what suggestions would you have for him or her? (Monitoring Understanding)
9. What was the most confusing aspect of the passage? What did you do when you were confused?
10. Do you think that this reading selection was too difficult, too easy, or just right for your reading ability? Why do you think that?

3–5 in Table 3). Then, after students read the second half of the passage, there are more questions related to analysis and synthesis, critical reading, and monitoring of understanding (questions 6–8). Finally, there are two self-assessment questions (questions 9 and 10).

### **The IRIS follow-up interview form**

Because a written test is not necessarily an accurate measure of each student's strategy use (e.g., some students struggle with writing or writing in English as their second language), we developed a follow-up oral interview to ask about students' answers to the pre-, during-, and postreading questions. Using the teacher record-keeping form of the IRIS, teachers copy answers given by students as they are probed about their written answers. Students may refer to their written tests that are in front of them. Ultimately, teachers can choose how many students with whom they would like to do follow-up interviews. Teachers might choose to interview only struggling students or students about whom they need more information.

In our first administration of the IRIS, we interviewed four proficient students, four average students, and four struggling students in each grade. We have found that students typically increase their comprehension strategy score by 1 or 2 points (out of 24 total points) on the oral interview form of the assessment. Because the scores of the written version are highly correlated with the oral version (.98), it is not necessary to do this with all students—just a subset of students about whom teachers would like to learn more or who, for various reasons, find writing answers more challenging than providing verbal responses.

Students are also asked to read aloud a short excerpt from the same reading passage. This is scored for accuracy and for sense (Rhodes & Shanklin, 1990). They are also asked to retell what they remember from their oral reading of the short excerpt. The back page of the IRIS provides a rubric for scoring the written (student) and oral (teacher record keeping) forms of the strategies assessment (see Table 4).

### **Administering and scoring the IRIS**

Teachers can give this written version of the IRIS to all of their students in approximately 30

minutes, and they can score it with the rubric provided on the back of the teacher record-keeping form. All participating schools gave this assessment to their students in grades 4 to 8 in autumn of 2003 and spring 2004 (there are two versions of the assessment, A and B). The project team then conducted follow-up, one-on-one interviews with 12 students per grade per school, because at that point not all teachers had been to the training sessions and did not feel they would have time to interview the students. (In the third year of the project, teachers conducted the follow-up interviews, which take about 15 minutes per student). Interrater reliability on 5% of the IRIS written and follow-up interviews ranges between 80 and 100% across the six strategies assessed.

We have found that the IRIS is highly correlated with the CAT subtests in reading (e.g., 0.79 correlation for grade 4 IRIS scores and grade 4 CAT Total Reading scores,  $N = 42$ ). This assures us that we are measuring similar constructs (reading comprehension ability), although we argue that the IRIS is more directly related to the particular goals of this project, which focus on reading strategy use and instructional decision making.

### **Emerging profiles of intermediate readers based on the IRIS**

From our experience conducting over 400 written and follow-up interviews, we have developed emerging profiles (cf. Wade, 1990) of intermediate readers that have been useful to teachers in understanding the strengths and instructional needs of their students in terms of the six reading comprehension strategies (and the oral reading and retelling measure). We include the caveat that while these profiles are useful for identifying general types of student readers, each student is unique in the combinations of strengths and weaknesses he or she brings to the reading process and may fit in more than one category. Also, this is just one assessment based on one nonfiction passage, and each student is interviewed in an unfamiliar situation. We recognize that a student may have a somewhat different profile in a different context. The following are the profiles we found.

**Word-level, processing problems.** Students with this profile are constructing little to no meaning

**TABLE 4**  
**Scoring rubric for written and oral responses**

	1	2	3	4
<b>Making Connections</b> Questions 1 & 2	No connections between background knowledge or experiences and the text in terms of purposes for reading or text ideas, genre, or writer's craft.	Makes some minimally related connections between background knowledge or experiences and the text in terms of purposes for reading, text ideas, genre, or writer's craft to support comprehension.	Makes appropriate connections between background knowledge or experiences and the text in terms of purposes for reading, text ideas, genre, or writer's craft that seem to support comprehension.	Makes rich connections between background knowledge or experiences and the text in terms of purposes for reading, ideas, genre, or writer's craft, including some that extend beyond the scope of the text.
<b>Active Meaning Construction</b> Question 3	No questioning, hypothesizing, or predicting; no connecting of ideas, inferencing, or use of text structure cues to support comprehension.	Does some minimally related questioning, hypothesizing, or predicting; or minimal connecting of ideas, inferencing, or use of text structure cues to support comprehension.	Appropriate use of questioning, hypothesizing, or predicting; or appropriate connecting of ideas, inferencing, or use of text structure cues to support comprehension.	Rich use of questioning, hypothesizing, or predicting; or appropriate connecting of ideas, inferencing, or use of text structure cues, including some that extend beyond the scope of the text.
<b>Engaging With the Text</b> Questions 4 & 5	No visualizing, engaging in world of text, perspective taking, or responding to author's craft.	Does some minimally related visualizing, engaging in world of text, perspective taking, or responding to author's craft to support comprehension.	Appropriate use of visualizing, engaging in world of text, perspective taking, or responding to author's craft to support comprehension.	Rich use of visualizing, engaging in world of text, perspective taking, or responding to author's craft to support and extend comprehension.
<b>Analysis and Synthesis</b> Question 6	No distinguishing more important ideas or events, drawing conclusions, identifying main ideas, themes, or connecting ideas to other texts.	Some minimal distinguishing of more important ideas or events, drawing of conclusions, main ideas, theme identification, or connecting of ideas to other texts to support comprehension.	Distinguishes more important ideas or events, draws conclusions, identifies main ideas, themes, or connects ideas to other texts to support comprehension.	Distinguishes important ideas or events, draws conclusions, identifies main ideas, themes, or connects ideas to other texts in rich ways that support and extend comprehension.
<b>Critical Reading</b> Question 7	No questioning of author, relating of text to self or world, evaluating ideas, or developing alternate interpretations.	Some minimally related questioning of author, relating of text to self or world, evaluating ideas, or developing alternate interpretations to support comprehension.	Appropriate use of questioning of author, relating of text to self or world, evaluating ideas, or developing alternate interpretations to support comprehension.	Rich use of questioning of author, relating of text to self or world, evaluating ideas, or developing alternate interpretations to support and extend comprehension.
<b>Monitoring Understanding</b> Questions 8, 9, & 10	No awareness of lack of understanding, revising hypotheses predictions, cross-checking of ideas, or using fix-up strategies.	Some awareness of lack of understanding but little or no revising of hypotheses, predictions, cross-checking of ideas, or use of fix-up strategies.	Awareness of lack of understanding and appropriate revising of hypotheses or predictions, cross-checking of ideas, or use of fix-up strategies to support comprehension.	Awareness of lack of understanding and rich use of revising of hypotheses or predictions, cross-checking of ideas, or use of fix-up strategies to support and extend comprehension.

Score for written responses

Questions 1 & 2	Making Connections	___/ 4
Question 3	Active Meaning Construction:	___/ 4
Questions 4 & 5	Engaging With the Text:	___/ 4
Question 6	Analysis and Synthesis:	___/ 4
Question 7	Critical Reading:	___/ 4
Questions 8, 9, & 10	Monitoring Understanding:	___/ 4
	Overall written response score:	___/24

Score for responses to oral interview

Making Connections	___/ 4
Active Meaning Construction:	___/ 4
Engaging With the Text:	___/ 4
Analysis and Synthesis:	___/ 4
Critical Reading:	___/ 4
Monitoring Understanding:	___/ 4
Overall oral response score:	___/24

from the passage. When we perform the oral reading measure, we find that they have decoding, word-identification, or fluency problems. These students need support for these skills in the context of meaningful text and discussions that support comprehension.

**“Reading” but not making meaning.** Students with this profile read fluently with perfect or near-perfect decoding and word identification but have little to no comprehension of the passage. They sometimes answer questions by “borrowing” phrases from boxed information, bolded information, or titles, but they do not provide evidence of constructing their own meaning from the text. These students need to begin using strategies that help them make meaningful connections to the text.

**Local meaning makers.** These students decode adequately, usually at the instructional level, and they retell fairly well but focus on meaning at the local or sentence-by-sentence level instead of constructing overall passage meaning or extending meaning beyond the passage. These students would benefit from strategies that help them construct meaning across the text, analyze and synthesize the ideas in the text, and read critically.

**Global meaning makers.** These students have rich connecting and engaging strategies and construct global meaning, but they only minimally attend to the structure and particulars of the passage. They may have some word-identification weakness, and they compensate by using a rich background knowledge and “guessing” at some of the content and structure. These students need to focus more on the structure and ideas in the text as evidence for their conclusions.

**The strategic majority.** Students who indicate no use or minimal use of some reading comprehension strategies and appropriate use of others (e.g., an overall average score of 14 to 16 out of 24). They often are able to connect ideas, engage with the text, and construct meaning (with attention to text as opposed to top-down readers), yet they may not read critically, make text-to-world connections, or monitor themselves when they are experiencing comprehension problems.

**The critics.** We hope all students become this type of reader. These students are very competent readers and exhibit strengths in most or all areas of comprehension and critical reading and thinking strategy use. They sometimes point out weaknesses in a reading passage or in a teacher’s questions.

## Making instructional decisions

When the project team shared the profiles with the teacher leaders from each school at a meeting in late autumn of 2003 (year 2), we noted that they recognized the types of students described and saw the potential usefulness of these profiles for making instructional decisions. For instance, at one school, teachers who scored their IRISs said it helped them to see the specific differences between their stronger and weaker readers in terms of the strategies the students were using. They decided to focus on constructing meaning and critical reading strategies because of the students’ lack of questioning of the text and the author. They decided to use a QAR (Question–Answer Relationship) Strategy (Raphael, 1982), noting that “when the question and answer are farther away from each other in the text it is more difficult for [students].”

Other teachers have chosen to use a mix of print and nonprint strategies to encourage the use of particular strategies, such as engaging (visualizing) and constructing meaning (predicting) for both literary and informational texts. In response, the mentors adapted a technique they called Read–Sketch–Predict to encourage students to read one stanza of a poem, sketch what they see, and then predict what will come next. After doing this for each stanza, the students wrote what they noticed about their thinking, which reinforces the Monitoring Understanding strategy. Many other examples have either been provided to the teachers in the resource tub or workshops or have been developed by the teachers at each school, including lessons that encourage the use of all six comprehension strategies with one informational book. For instance, along with the book *Should There Be Zoos?* (Stead, 2002), teachers created the following strategy lessons: K-W-L sheets (Ogle, 1986) for drawing on background knowledge, using illustrations to predict, collaboratively filling out sheets that provide a column for key opinions



and evidence in the text, developing arguments for and against zoos, writing down arguments, and revising papers with a buddy using guidelines for persuasive writing.

It is important to note that the IRIS was developed to complement and inform the kinds of strategy work teachers were already doing, but it also gives them an opportunity to see if students are developing their use of particular strategies after instructional implementation. Some schools have chosen to use this measure as one indicator of literacy growth in their yearly school accountability contracts. Such a use illustrates our notion of a “situated” assessment that integrates assessment and practice, informing teachers rather than simply holding them accountable with measures unrelated to their practices.

## Support and encourage accountability and literacy improvement

When we began collaborating on this project in the spring of 2002, we did not envision creating a specific assessment tool of reading comprehension strategies to support or evaluate the project. As the project developed, and we began to see the need for another kind of measure, we looked at a range of commercial informal reading inventories and realized none of them fully matched the goals of our project or the needs of the teachers. Over time we have simultaneously developed the instrument and conceptualized the role of assessment in this large professional development and reform project. Our on-the-ground, collaborative theorizing has led us to believe that not only do measures of projects such as this need to be valid in their content but also in terms of consequences and uses of the assessment. Teachers, school board consultants, and researchers were all actively engaged in developing the IRIS as it went through many draft stages, constantly being refined in consultation with stakeholders and as a result of their feedback. The transactional validity is reflected in this collaborative development and piloting of the assessment, as well as in the careful consideration of its uses. The uses (and consequences) of the instrument are consistent with the layers of systemic goals among the group using it.

Teachers wanted a tool that they could use and score themselves—one that would inform their instruction and their students’ understanding of what kinds of strategies were important in the reading process. In this way, the instructional changes “engendered by the use of the test” have the potential to “contribute to the development of knowledge and/or skills that the test purportedly measures” (Frederiksen & Collins, 1989, p. 27).

We know that the assessment is limited to one passage (per grade level) that any individual student may or may not find interesting. It is therefore a measure that should be used in relation to other classroom assessments to gauge a student’s use of reading comprehension strategies. Over time, we will have even more information about how teachers use the assessment as part of their instructional decision making in relation to teaching reading comprehension strategies in their classrooms and to what extent reading strategy use increases among their students. We also will have more information about how they might use such an assessment for accountability. For instance, one inner-city secondary school analyzed its autumn and spring scores and calculated how many students were at each of the 4 levels and set goals for the next year: 100% of students scoring at least at the minimal range in using the six comprehension strategies (with overall scores of 12 and above out of 24) and 75% of students using appropriate comprehension strategies (scores of 18 and above out of 24).

In this way, we hope to respond to the growing demand for assessment approaches that are sensitive to the contexts in which they are used, build on notions of consequential and transactive validity, and allow for critical inquiry into the relationship between curriculum and instruction. This approach also supports and encourages teachers and school administrators who set their own goals for accountability and improvement in literacy.

For more information about the IRIS, please e-mail the first author at [theresa.rogers@ubc.ca](mailto:theresa.rogers@ubc.ca).

**Rogers teaches at the University of British Columbia (Department of Language and Literacy Education, Ponderosa E, 2034 Lower Mall, Vancouver, BC V6T 1Z4, Canada). Winters and Bryan are doctoral students at the same university. Price teaches at Elsie Roy Elementary School in Vancouver. McCormick,**

**House, and Mezzarobba are members of the Vancouver School Board. Sinclair is a faculty associate at Simon Fraser University in Burnaby, British Columbia.**

## References

- Anders, P.L., Hoffman, J.V., & Duffy, G.G. (2000). Teaching teachers to teach reading. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 719-742). Mahwah, NJ: Erlbaum.
- Cassidy, J., & Cassidy, D. (2004/2005, December/January). What's hot, what's not for 2005. *Reading Today*, 22(3), pp. 1, 8-9.
- Dillon, D. (2003). In leaving no child behind have we forsaken individual learners, teachers, schools and communities? Presidential address. In C.M. Fairbanks, J. Worthy, B. Maloch, J.V. Hoffman, & D.L. Schallert (Eds.), *52nd Yearbook of the National Reading Conference* (pp. 1-31). Oak Creek, WI: National Reading Conference.
- Duke, N.K., & Pearson, P.D. (2002). Effective practices for developing reading comprehension. In A.E. Farstrup & S.J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 205-242). Newark, DE: International Reading Association.
- Durkin, D. (1978/1979). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly*, 14, 481-533.
- Fielding, L.G., & Pearson, P.D. (1994). Synthesis of research/reading comprehension: What works. *Educational Leadership*, 51(5), 62-68.
- Fullan, M. (1992). *Successful school improvement*. Buckingham, England: Open University Press.
- Frederiksen, J.R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, 18(9), 27-32.
- Hill, P. (1998). Shaking the foundations: Research-driven school reform. *School Effectiveness and School Improvement*, 9, 419-436.
- Hubley, A.M., & Zumbo, B.D. (1996). A dialectic on validity: Where we have been and where we are going. *The Journal of General Psychology*, 123, 207-215.
- International Reading Association & National Council of Teachers of English. (1994). *Standards for the assessment of reading and writing*. Newark, DE; Urbana, IL: Authors.
- Johnston, P. (1998). The consequences of the use of standardized tests. In S. Murphy (Ed.), *Fragile evidence: A critique of reading assessment* (pp. 89-101). Mahwah, NJ: Erlbaum.
- Langer, J. (1999). *Excellence in English in middle and high school: How teachers' professional lives support student achievement*. Albany: Center on English Learning and Achievement, State University of New York.
- Luke, A., & Elkins, J. (1998). Editorial: Reinventing literacy in "New Times." *Journal of Adolescent & Adult Literacy*, 42, 4-7.
- Moss, P. (1998). Enlarging the dialogue in educational measurement. Voices from interpretive traditions. *Educational Researcher*, 25(1), 20-28.
- Murphy, S. (1998). *Fragile evidence: A critique of reading assessment*. Mahwah, NJ: Erlbaum.
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- Ogle, D. (1986). K-W-L: A teaching model that develops active reading of expository text. *The Reading Teacher*, 39, 564-570.
- Pearson, P.D., & Dole, J.A. (1987). Explicit comprehension instruction: A review of research and a new conceptualization of instruction. *The Elementary School Journal*, 88, 151-165.
- Pearson, P.D., Roehler, L.R., Dole, J.A., & Duffy, G.G. (1992). Developing expertise in reading comprehension. In S.J. Samuels & A.E. Farstrup (Eds.), *What research has to say about reading instruction* (2nd ed., pp. 145-199). Newark, DE: International Reading Association.
- Raphael, T.E. (1982). Question-answering strategies for children. *The Reading Teacher*, 36, 186-191.
- Rhodes, L.K., & Shanklin, N.L. (1990). Miscue analysis in the classroom. *The Reading Teacher*, 44, 252-254.
- Stead, T. (2002). *Should there be zoos? A persuasive text*. New York: Mondo.
- Taylor, B.M., Pearson, P.D., Clark, K., & Walpole, S. (2002). Effective schools and accomplished teachers: Lessons about primary-grade reading instruction in low-income schools. In B.M. Taylor & P.D. Pearson (Eds.), *Teaching reading: Effective schools, accomplished teachers* (pp. 3-72). Mahwah, NJ: Erlbaum.
- Tierney, R.J., Crumpler, T.P., Bertelsen, C.D., & Bond, E.L. (2003). *Interactive assessment: Teachers, parents and students as partners*. Norwood, MA: Christopher-Gordon.
- Wade, S.E. (1990). Using think-alouds to assess comprehension. *The Reading Teacher*, 43, 442-451.